(12) UK Patent Application (19) GB (11) 2 359 457 (13) A

(43) Date of A Publication 22.08.2001

(21) Application No 0003897.6

(22) Date of Filing 18.02.2000

(51) INT CL[7]
H04M 1/27 , H04Q 7/32

(52) UK CL (Edition S)
H4L LEUF L1H10

(56) Documents Cited
GB 2344028 A GB 2327554 A EP 0739121 A2

(58) Field of Search
UK CL (Edition R) H4K KBNJ KFH , H4L LEUF
INT CL[7] H04M 1/27 , H04Q 7/32
ONLINE: WPI, EDOC, JAPIO

(71) Applicant(s)
Nokia Mobile Phones Limited
(Incorporated in Finland)
Keilalahdentie 4, 02150 Espoo, Finland

(72) Inventor(s)
Allan Suonperä
Juha Kurkilahti
Elise Levanto
Peter Ollikainen

(74) Agent and/or Address for Service
Nokia IPR Department
Nokia House, Summit Avenue, Southwood,
FARNBOROUGH, Hampshire, GU14 0NG,
United Kingdom

(54) Abstract Title
Hand portable phone supporting voice-controlled hands-free operation

(57) A hand portable phone for communication via a communication network, includes means 20 for establishing and terminating a call over said network. The terminal has a first microphone 6 and a first speaker 5 and acting as user interface during a call and, means for supporting speech control of the operations performed in the terminal. An accessory means 25 that, when connected to the terminal, acts as an audio user interface during a hands free call includes a second microphone 21 and a second speaker 22, and means 20 for activating the speech control means and for handling the call. Speech control is activated for a predetermined period after the means 20 is actuated in the idle mode of the phone.

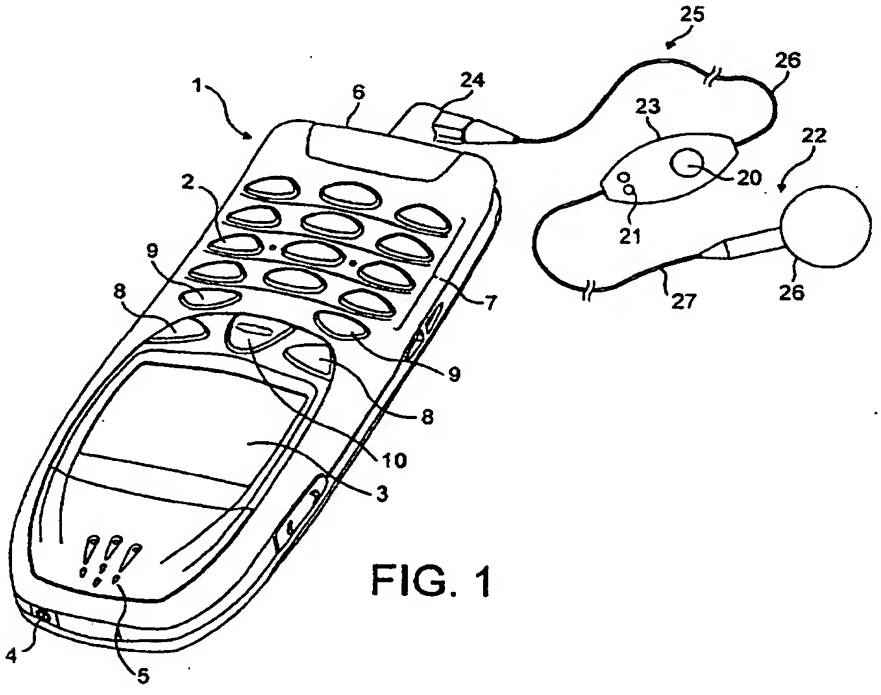

FIG. 1

GB 2 359 457 A

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

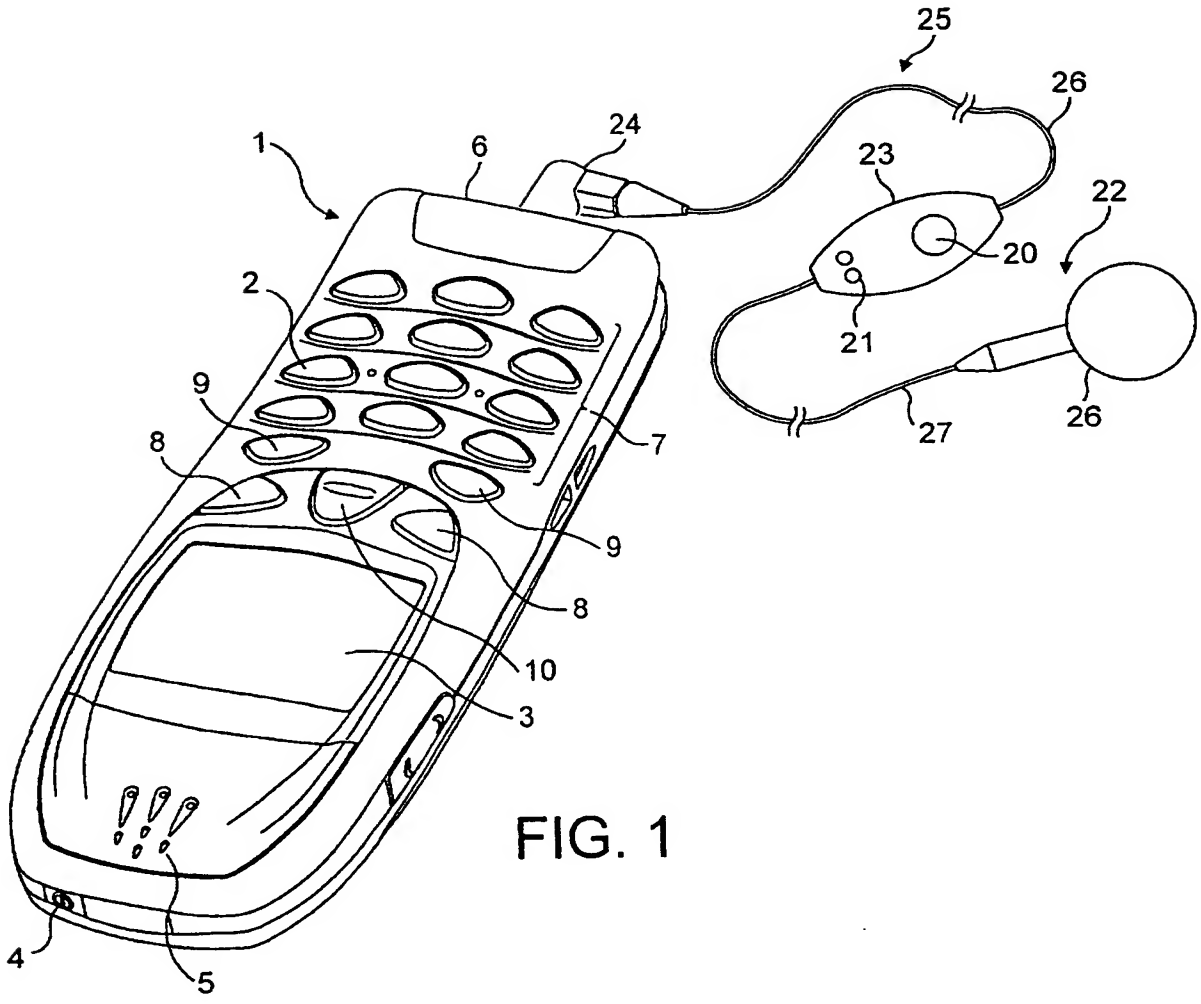

FIG. 1

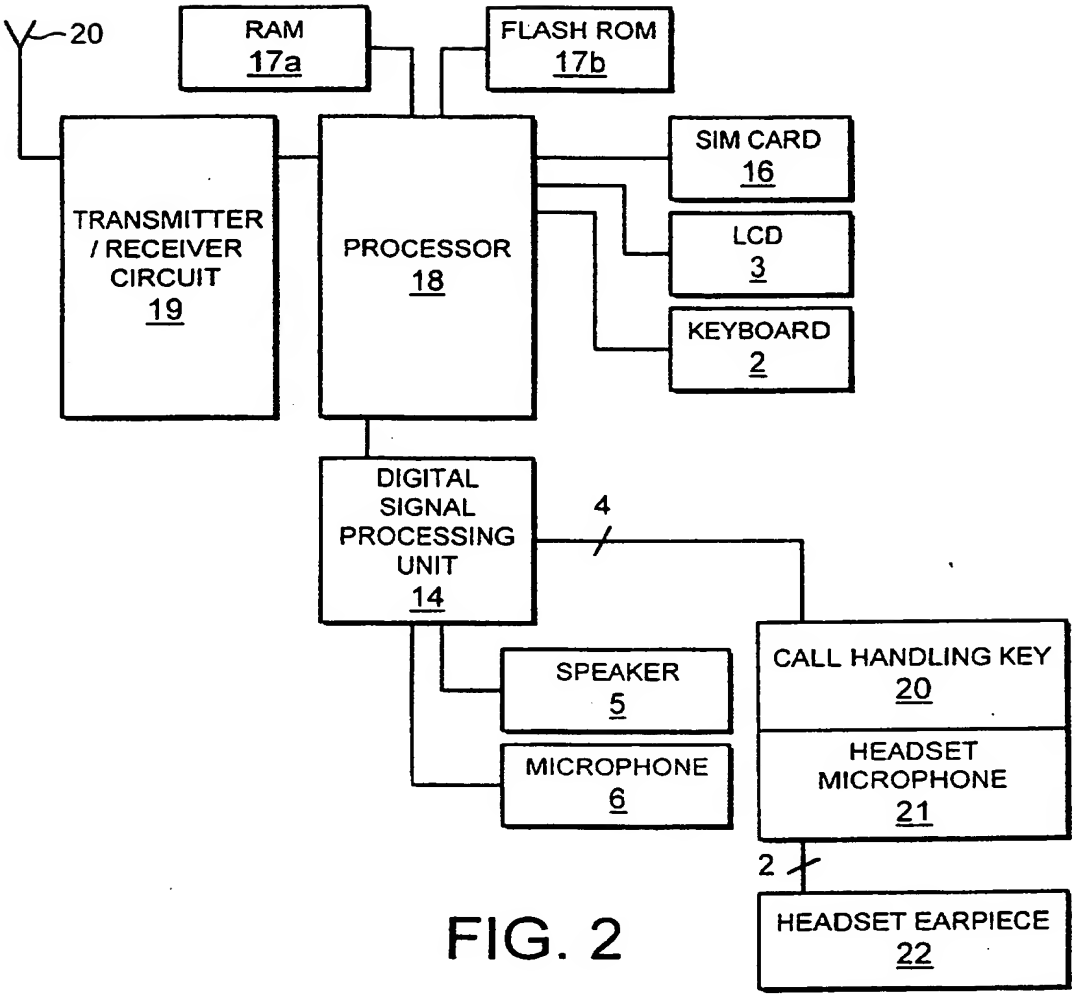
20
RAM
17a
FLASH ROM
17b
TRANSMITTER / RECEIVER CIRCUIT
19
PROCESSOR
18
SIM CARD
16
LCD
3
KEYBOARD
2
DIGITAL SIGNAL PROCESSING UNIT
14
4
SPEAKER
5
MICROPHONE
6
CALL HANDLING KEY
20
HEADSET MICROPHONE
21
2
HEADSET EARPIECE
22

FIG. 2

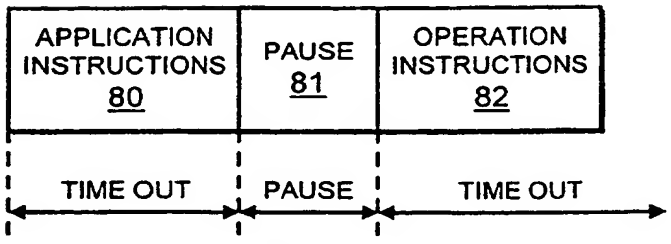
APPLICATION INSTRUCTIONS
80
PAUSE
81
OPERATION INSTRUCTIONS
82
TIME OUT
PAUSE
TIME OUT

FIG.5

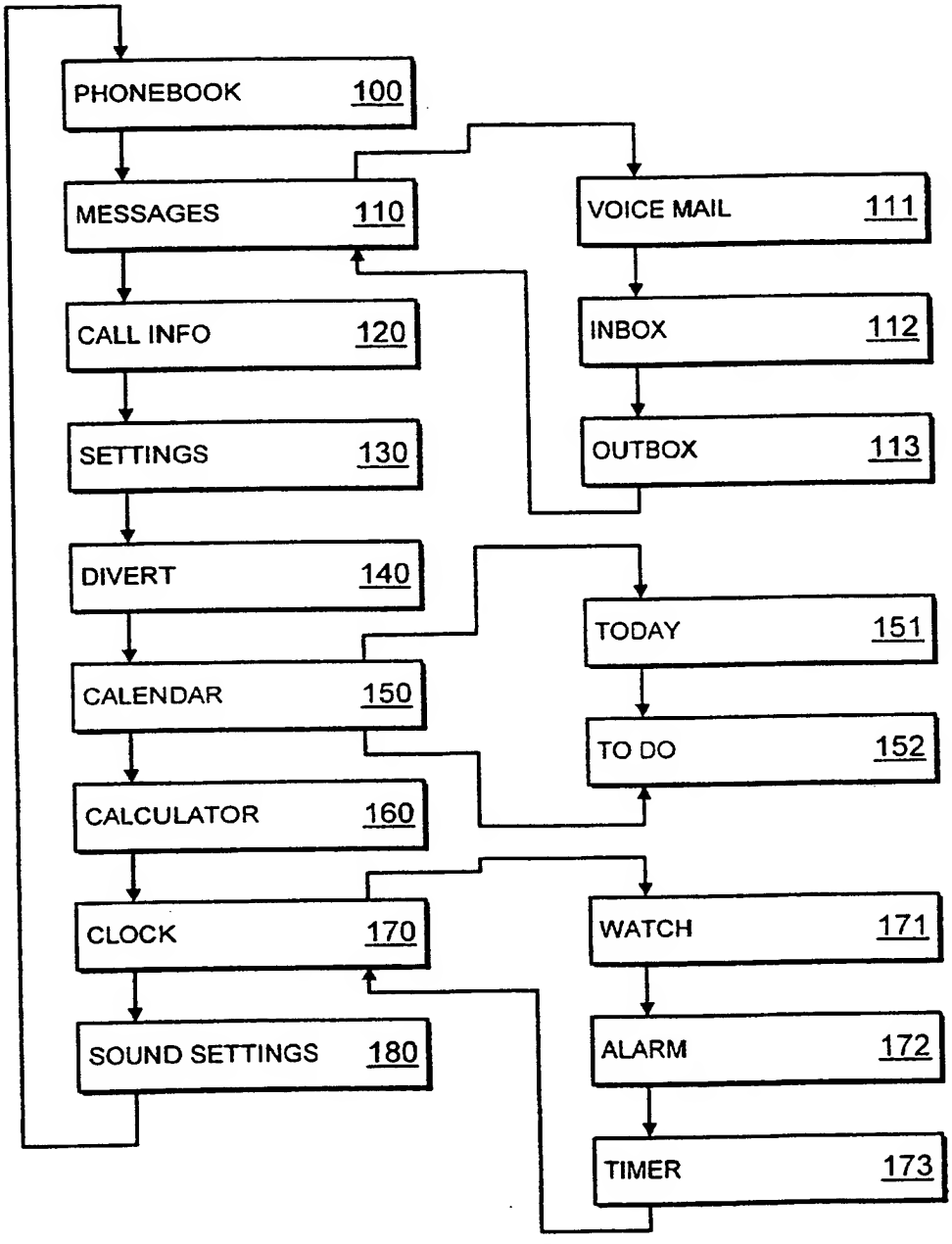
PHONEBOOK 100
MESSAGES 110
VOICE MAIL 111
CALL INFO 120
INBOX 112
SETTINGS 130
OUTBOX 113
DIVERT 140
TODAY 151
CALENDAR 150
TO DO 152
CALCULATOR 160
CLOCK 170
WATCH 171
SOUND SETTINGS 180
ALARM 172
TIMER 173

FIG. 3

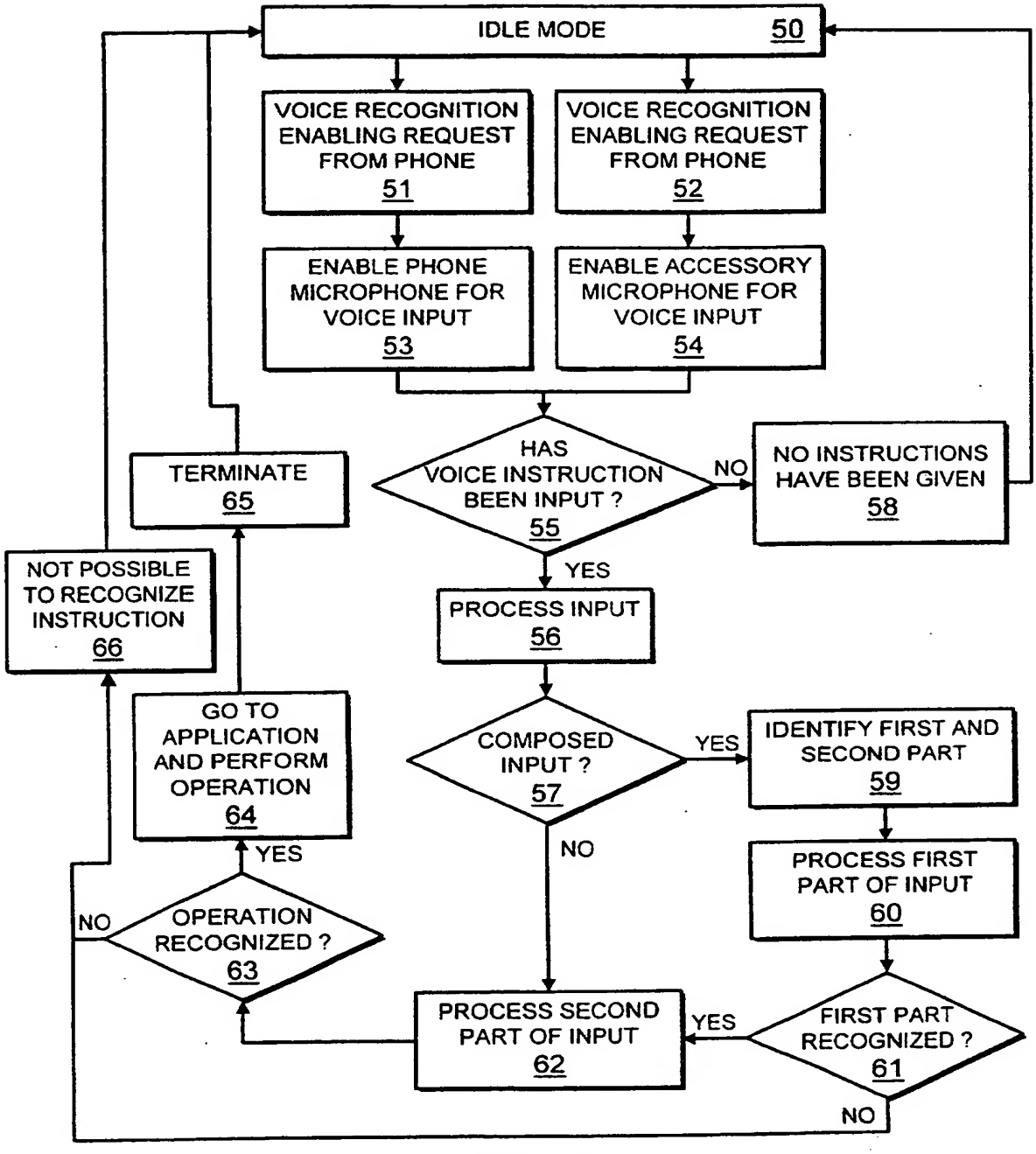
IDLE MODE 50
VOICE RECOGNITION ENABLING REQUEST FROM PHONE 51
VOICE RECOGNITION ENABLING REQUEST FROM PHONE 52
ENABLE PHONE MICROPHONE FOR VOICE INPUT 53
ENABLE ACCESSORY MICROPHONE FOR VOICE INPUT 54
HAS VOICE INSTRUCTION BEEN INPUT ? 55
NO
NO INSTRUCTIONS HAVE BEEN GIVEN 58
YES
PROCESS INPUT 56
COMPOSED INPUT ? 57
YES
IDENTIFY FIRST AND SECOND PART 59
NO
PROCESS FIRST PART OF INPUT 60
FIRST PART RECOGNIZED ? 61
YES
NO
PROCESS SECOND PART OF INPUT 62
OPERATION RECOGNIZED ? 63
NO
YES
GO TO APPLICATION AND PERFORM OPERATION 64
NOT POSSIBLE TO RECOGNIZE INSTRUCTION 66
TERMINATE 65

FIG. 4

A hand portable phone supporting speech control of its operation.

The invention relates to a method of controlling a hand portable phone having speech control means for supporting speech control of its operation.

The applicant launched recently a new GSM dual band phone named Nokia 8850™ having speech control means for supporting speech control of its operation.

During recent years it has become widely used to integrate speech control for establishing calls. However until now this feature has more or less been used to prove the technical level of the manufacturer instead of giving the user real benefit.

An object of the invention is to provide a method of controlling a hand portable phone having speech control means for supporting speech control of its operation for improving the hands free operation of the phone.

According to the invention there is provided a method of controlling a handportable phone having speech control means for supporting speech control of its operation, said phone has an accessory means, that, when connected to the terminal, acts as a user interface during a hands free call, said accessory means includes a call handling key for accepting an incoming call and for terminating a call for activating the speech control means wherein the method includes step of activate the speech control means for a predetermined period when said call handling key is activated in idle mode of the phone.

In many countries you are only allowed to have telephone conversions, as long as this is based on a so-called hands free concept. This means that the user must have an earpiece placed in this ear. The earpiece is according to

the preferred embodiment connected to the telephone by means of a wire connection also carrying a microphone and a call-handling key for accepting an incoming call and for terminating calls. According to the invention this accessory also includes means for enabling the speech control of the operation of the telephone.

By integrating this functionality in a single multi-functionality key, the user can easily control the telephone in a non-ambiguous way – this risk of making mistakes is eliminated. According to prior art the user has to enable the speech control by pressing one of at least fifteen keys on the telephone itself. In other words the invention offers the user an improved speech control concept.

Preferably the speech control means will be enabled to receive speech instructions during a predetermined period of time once the call-handling key is activated.

According to a further aspect of the invention, there is provided a hand portable telephone for communication via a communication network. The telephone includes means for establishing and terminating a call over said network, first audio means integrated into the terminal as user interface during a call, said first audio means includes a microphone and a speaker, and means for supporting speech control of the operations performed in the telephone. The telephone does furthermore include an accessory means, that, when connected to the terminal, acts as a user interface during a hands free call, said accessory means includes second audio means that includes a microphone and a speaker, and said accessory means includes means for activating the speech control means.

Preferably the accessory means is a headset connected to the phone by means of a wire and a connector plug, and said means for activating the

speech control means is integrated into a key for accepting an incoming call. However the accessory means could be a headset connected via a wireless connection, e.g. based on the Bluetooth protocol, to the telephone.

Preferably when activating said key for accepting an incoming call in idle mode of the phone, the speech control means will become activated for a predetermined period.

According to the preferred embodiment of the invention the activation of the speech control means in said accessory means will cause the means for supporting speech control to use the microphone of the second audio means in the accessory means as input means for the speech instructions.

According to the preferred embodiment of the invention the speech control means will based on speech instructions be able to find and call phone numbers from a phone number database stored in a phone by recognizing the speech instructions as a pre-stored voice tag stored in association with the phone number to be called.

According to an alternative embodiment of the invention the speech control means will based on speech instructions be able to perform operations from the phone menu structure by recognizing the speech instructions as a pre-stored voice tag stored in association with the operation to be performed.

According to a second alternative embodiment of the invention the speech control means will, when processing the speech instructions, firstly look for a speech sequence identifying an application available in the telephone, and secondly look for a speech sequence identifying the operation in said application to be performed. By saying "calendar" followed by a pause, e.g. with duration of 0.8 seconds and then "to-do", the telephone could then play

back recorded voice tag stored in the to-do list of the calendar or hear synthesized speech based on the written input in this list.

For a better understanding of the present invention and to understand how the same may be brought into effect reference will now be made, by way of example only, to accompanying drawings, in which:-

Fig. 1 illustrates in perspective a preferred embodiment of a hand portable phone with an accessory according to the invention.

Fig. 2 schematically shows the essential parts of a telephone for communication with a cellular or cordless network.

Fig. 3 schematically shows an exemplary menu structure of a phone according to the invention.

Fig. 4 shows a flow diagram illustrating the concept according to the invention for speech controlling the operations of a hand portable telephone.

Fig. 5 shows a composite voice instruction according to the invention.

Fig. 1 shows a preferred embodiment of a phone according to the invention, and it will be seen that the phone, which is generally designated by 1, comprises a user interface having a keypad 2, a display 3, an on/off button 4, a speaker 5, and a microphone 6 (openings present in the bottom of the phone and therefore not visible in the present view). The phone 1 according to the preferred embodiment is adapted for communication via a cellular network, e.g. the GSM 900/1800 MHz network.

According to the preferred embodiment the keypad 2 has a first group 7 of keys as alphanumeric keys, two soft keys 8, two call handling keys 9, and a

cursor navigation key 10. The present functionality of the soft keys 8 is shown in separate fields in the display 3 just above the keys 8, and the call handling keys 9 are used for establishing a call or a conference call, terminating a call or rejecting an incoming call.

A headset 25 is plugged into the phone 1 via the bottom connector of the phone (not shown, but corresponds to the connector used in handsets Nokia 5110 and Nokia 6110) by means of a Jack plug 24. The headset 25 furthermore includes an earpiece part 22 having a speaker enclosed in a foam rubber pad 26 for being placed in outer part of auditory means of the ear of the user, and microphone body 23 including a microphone 21 (holes are shown) and a call handling key 20. Two cords 26 and 27 connect the individual part of the headset 25. The cord 26 is a four wire cord – one common ground wire, one microphone wire, one speaker wire and one signaling wire.

Fig. 2 schematically shows the most important parts of a preferred embodiment of the phone, said parts being essential to the understanding of the invention. The processor 18 controls the communication with the network via the transmitter/receiver circuit 19 and an internal antenna 20.

The microphone 6 transforms the user's speech into analog signals, the analog signals formed thereby are A/D converted in an A/D converter (not shown) before the speech is encoded in a digital signal processing unit 14 (DSP). The encoded speech signal is transferred to the processor 18, which i.a. supports the GSM terminal software. The processor 18 also forms the interface to the peripheral units of the apparatus, including a RAM memory 17a and a Flash ROM memory 17b, a SIM card 16, the display 3 and the keypad 2 (as well as data, power supply, etc.). The digital signal-processing unit 14 speech-decodes the signal, which is transferred from the processor 18 to the earpiece 5 via a D/A converter (not shown).

When the headset 25 is plugged into the phone the speaker 5 and the microphone 6 will become disabled. Instead the speaker 22 and the microphone 21 will act as audio interface between the user and the digital signal-processing unit 14.

The invention is not specifically related to the voice recognition as such and therefor this will not be described in detail. However US 5,481,595 describes an example of such a voice control concept.

It is widely used to integrate voice recognition algorithms into a phone for making phone calls. This is basically for reducing the required processor power needed for processing the inputted voice instructions and for comparing these instructions with pre-stored voice tags. When the comparison is successful the processor 18 of the phone automatically executes the operation identified by means of the inputted voice instruction.

According to the invention the phone supports a composite voice instruction. This means that the user firstly identifies the application and then the operation in the voice instruction. Preferably the two sub-instructions are separated by a short pause, e.g. with duration of 0.8 seconds. By pressing the call handling key once, and then saying "messages", followed by a short break and then continue by saying "voice mail", the phone will dial up the voice mail box service if these voice tags has been stored in the phone.

With reference to fig. 4 it is seen that the phone 1, when it is in idle mode (step 50), may receive a request for enabling voice recognition – e.g. if the user has pressed the navigation key 10 (step 51) or the user has pressed the call handling key 20 on the headset (step 52). Then the phone 1 enables either the phone microphone 6 (step 53) or the headset microphone 21 (step 54) for receiving a voice input.

In step 55 it is checked whether a voice instruction has been received or not. If no instruction has been given it is decided in step 58 that no instructions has been given, and the phone goes back to idle mode again.

If an instruction has been given the digital signal processing unit 14 starts to process the voice input in step 56, and in step 57 it is evaluated whether the input is a composite input or not. If the input in step 57 is deemed to be composite, the algorithm identifies the first and the second part of the input in step 59, and the first part is processed in step 60 in order to find matching pre-stored voice tags. If no matching is found for the first part in step 61, a response message is given in step 66 saying that it has not been possible to recognize the voice instruction. This message may be an alert tone accompanied with a text message in the display, play back of a recorded message or a synthesized speech.

If matching for the first part is found in step 61, the algorithm starts to process the second part of the voice instruction in step 62. This second part of the voice instruction may also be called the operation part because the operation to be performed is identified by the part. If the algorithm in step 57 deems the voice instruction to be a non-composite input – according to the preferred embodiment of the invention this means that the algorithm interprets the voice instruction as being a name in the phonebook to be called (phonebook is default). Then the algorithm starts to process the input in step 62. In step 63 it is evaluated whether is has been possible to find a matching pre-stored voice tag. If no match is found, a response message is given in step 66 saying that it has not been possible to recognize the voice instruction.

If a match has been found in step 63, an algorithm instructs the processor 18 to go to the identified application and perform the identified operation. When

the operation, e.g. a call, is terminated in step 65 the phone goes back to idle mode (step 50).

With reference to fig. 5 the algorithm starts to listen for a voice instruction for a first period (the first time out) of e.g. 3 seconds. If the algorithm does not detect something that might be interpreted as voice instruction within this first period of time the algorithm deems in step 55 that no voice message has been received. If a voice message has been detected the inputted voice is recorded for a period not exceeding a maximum of e.g. 15 seconds.

During this recording the algorithm looks for the termination of the voice message, and if a break occurs the algorithm investigates the duration of the break. If the break is less than, e.g. 0.5 seconds before the voice instruction is detected the algorithm deems this voice instruction as being a continuously part of the voice instruction before the break. If the break is longer, e.g. 0.5 seconds, the break is deemed to be a pause before the next voice instruction, that will be an independent sub-voice instruction in a composite voice instruction. If the break is longer, e.g. 2.0 seconds, the voice instruction is deemed to be terminated. If longer composite voice instructions are allowed the process of investigating the termination of the second, third and further sub-voice instruction could be done as explained for the first voice instruction.

Fig. 3 illustrates the menu structure of the phone according to the preferred embodiment of the invention. The menu has the following main menus "Phonebook" 100, "Messages" 110, "Call info" 120, "Setting" 130, "Divert" 140, "Calendar" 150, "Calculator" 160, "Clock" 170 and "Sound settings" 180. These menus have sub-menus, and the "Messages" menu 110 has e.g. three sub-menus "Voice mail" 111, "Inbox" 112 and "Outbox" 113.

When the user activates the voice recognition algorithm by pressing the call handling key 20, says "messages", holds a pause for 0.5-2.0 seconds, and

the says "voice mail", the phone then calls the voice mail box of the user. This requires that the user has set the voice tags at the menus and sub-menus.

According to an alternative embodiment of the invention the phone is provided with a program for generating a synthetic voice in response to a text message. Then user can access the calendar sub-menus "Today" 151 and "To-do" 152 and listen to the inputted activities translated from text to synthetic voice.

The user will furthermore have the possibility to use the calculation power of the phone by accessing the clock sub-menus "Watch" 171, "Alarm" 172 and "Timer" 173. Then the user may use the voice control to control the timer operations, e.g. it might be valuable to activate the timer as explained above, instructing the timer to "Start" and "Stop". The synthetic voice program could then be used for giving the result.

CLAIMS

1. A method of controlling a hand portable phone having speech control means for supporting speech control of its operation, said phone has an accessory means, that, when connected to the terminal, acts as an user interface during a hands free call, said accessory means includes a call handling key for accepting an incoming call and for terminating a call for activating the speech control means wherein the method includes step of activate the speech control means for a predetermined period when said call handling key is activated in idle mode of the phone.

2. A method according to claim 1, wherein said speech control means receives speech instructions during said predetermined period of time.

3. A method according to claim 2, wherein said speech control means as default uses the speech instructions for finding and calling phone numbers stored in a phone number database in a phone when the speech instructions matches a voice tag stored in association with the phone number to be called.

4. A method according to claim 2, wherein said speech control means firstly looks for a speech sequence identifying an application available in the phone, and secondly looks for a speech sequence identifying the operation to be performed.

5. A hand portable phone for communication via a communication network, said terminal includes:

- means for establishing and terminating a call over said network;
- first audio means integrated into the terminal as user interface during a call, said first audio means includes a microphone and a speaker;

- means for supporting speech control of the operations performed in the terminal;
- an accessory means, that, when connected to the terminal, acts as a user interface during a hands free call;
- said accessory means includes second audio means that includes a microphone and a speaker, and
- said accessory means includes means for activating the speech control means.

6. A hand portable phone according to claim 5, wherein said accessory means is a headset connected to the phone by means of a wire and a connector plug, and said means for activating the speech control means is integrated into a key for accepting an incoming call.

7. A hand portable phone according to claim 6, wherein activating said key for accepting an incoming call in idle mode of the phone will activate the speech control means for a predetermined period.

8. A hand portable phone according to claim 7, wherein activation the speech control means in said accessory means will cause the means for supporting speech control to use the microphone of the second audio means in the accessory means as input means for the speech instructions.

9. A hand portable phone according to claim 8, wherein said speech control means based speech instructions is able to find and call phone numbers from a phone number database stored in a phone by recognizing the speech instructions as an pre-stored voice tag stored in association with the phone number to be called.

10. A hand portable phone according to claim 8, wherein said speech control means based speech instructions is able to perform operations from the

phone menu structure by recognizing the speech instructions as a pre-stored voice tag stored in association with the operation to be performed.

11. A hand portable phone according to claim 8, wherein said speech control means based speech instructions is able to perform operations from applications included in the phone by recognizing the speech instructions as a pre-stored voice tag stored in association with the operation to be performed.

12. A hand portable phone according to claim 11, wherein said speech control means when processing the speech instructions firstly looks for a speech sequence identifying the application, and secondly looks for a speech sequence identifying the operation to be performed.

13. A hand portable phone substantially as herein before described with reference to the accompanying drawings, and/or as shown therein.

The Patent Office

13

INVESTOR IN PEOPLE

Application No: GB 0003897.6
Claims searched: all

Examiner: Nigel Hall
Date of search: 9 August 2000

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.R): H4L (LEUF); H4K (KBNJ, KFH)

Int Cl (Ed.7): H04M 1/27; H04Q 7/32

Other: Online: WPI, EDOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
X,E	GB 2344028 A	(NEC) See whole document	5 at least
X,Y	GB 2327554 A	(NOKIA) See whole document-	5 at least
Y	EP 0739121 A2	(TEXAS) See lines 25-29 of col.4	7

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.